

Financial Statistics in India: Current Contours and Future Challenges

Partha Ray*

Director, National Institute of Bank Management, Pune

Received: 14 August 2026; Revised: 01 September 2026; Accepted: 10 September 2026

Abstract

This paper examines the architecture and challenges of financial statistics in India, arguing that while finance now constitutes a major share of India's gross value added, the statistical measurement of the financial sector remains somewhat fragmented. The paper describes how India's financial data collection is deliberately distributed across sector-specific regulators—the RBI for banking, SEBI for capital markets, IRDAI for insurance, PFRDA for pensions, and others—each operating within distinct taxonomic frameworks and publication schedules. The paper identifies six critical challenges: (a) financial innovation, particularly fintech and cryptocurrency; (b) the rise of non-banking financial companies and informal lending; (c) data gaps on household balance sheets and micro, small, and medium enterprises; (d) implementation of the Digital Personal Data Protection Act of 2023; (e) increasing cross-border financial integration; (f) the emergence of green finance. Finally, artificial intelligence and alternative data sources present both opportunities and risks, including algorithmic bias and opacity. The paper concludes that robust financial statistics require greater inter-institutional coordination, modernized data governance frameworks, and systematic investment in statistical infrastructure to effectively measure an increasingly sophisticated financial system.

*This article is based on a talk delivered at the Session on Financial Statistics in the 28th Annual Conference of the Society of Statistics, Computers and Applications, on Innovative Trends in Statistical Science, Data Analytics and Interdisciplinary Research-2026 (ITSSDAIR-2026), co-hosted by the Department of Statistics and Centre for Advanced Studies, Savitribai Phule Pune University, on January 5, 2026. The author is indebted to Dr. R.B. Barman, Dr. A.C. Kulshreshtha, Dr. Ajit Joshi, and Dr. Praggya Das, for their comments on the talk. The usual disclaimer applies.

Key words: India; Financial Statistics; RBI; SEBI; Artificial Intelligence.

1. Introduction

The architecture of financial statistics in any country perhaps reflects both the institutional maturity and complexity of the financial sector. India is no exception. Any serious engagement with financial data in India, thus, must grapple with a fundamental premise: finance is not merely ancillary to the economy, but is core to economic activity. Interestingly, economists have diametrically opposite views on the issue of direction of causality between finance and growth. While British economist Joan Robinson famously said that “finance follows enterprise,” Nobel Laureate US economist Merton Miller argued that financial markets actively could spur real economic activity.

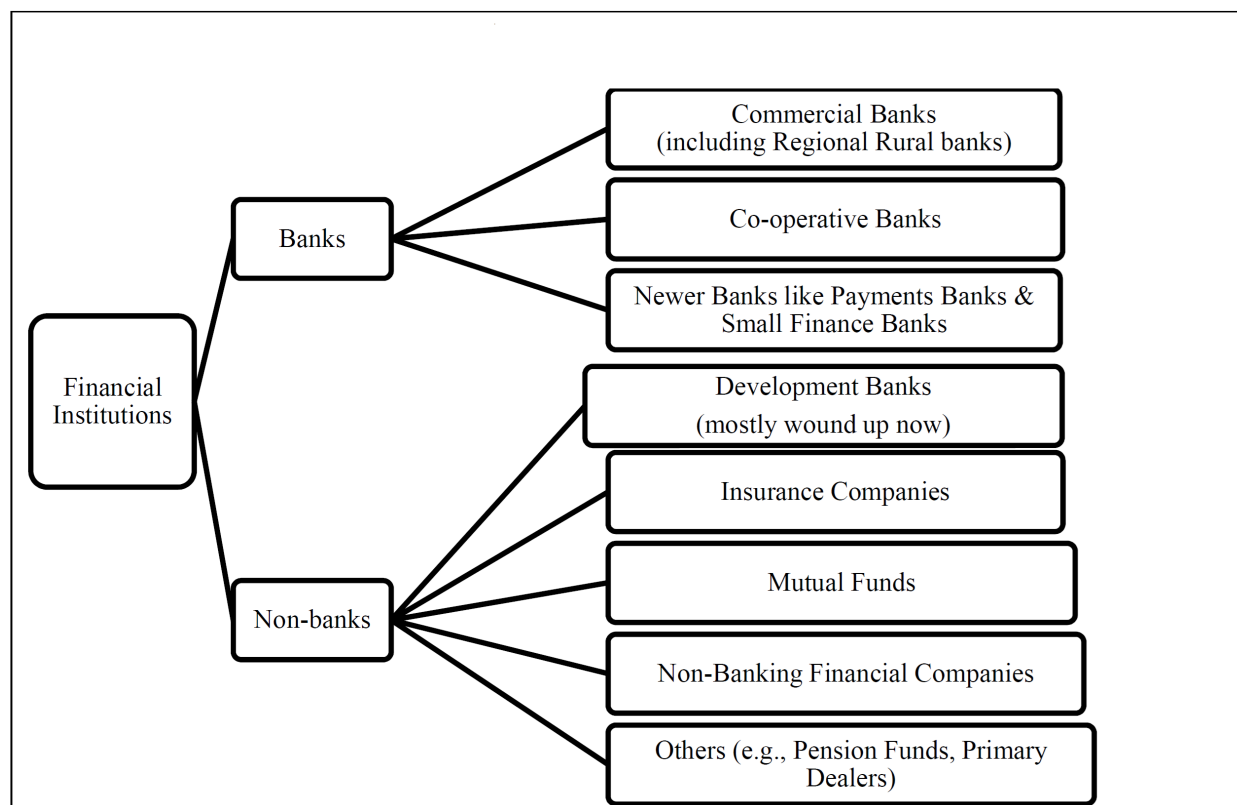
In terms of recent Indian specifics, the “financial, real estate, and professional services” sectors together constitute nearly 24 percent of India’s gross value added, a figure that has grown steadily from approximately 19 percent in 2011-12. This expansion underscores the importance of finance in the national economy - both as a component of output and as a critical infrastructure for monetary policy, financial stability, and resource allocation across the economy.

Has this prominence of finance in economic life has not been matched by corresponding clarity, consistency or excitement in the collection, standardization, and dissemination of financial statistics? Do the real sector statistics compilers have skepticism of their financial sector counterparts and vice versa? What are the challenges facing the financial statistician? The present essay delves into some of these questions.

2. The institutional framework

It may be apposite to begin with a description of the Indian financial system. Figure 1 depicts the major constituents of Indian financial system. India’s financial statistical system operates through a deliberately distributed architecture, in which responsibility for data collection and compilation rests with sector-specific regulators. The Reserve Bank of India (RBI) oversees banking statistics, extracting granular information from scheduled commercial banks on balance sheet composition, credit deployment, deposits, interest rates, and asset quality. It also compiles data on non-banking financial companies (NBFCs), an increasingly significant component of the financial system. The Securities and Exchange Board of India (SEBI), India’s stock market (comprising India’s two major stock exchanges, NSE and BSE) regulator, maintains statistics on equity and debt market turnover, market capitalization, IPO and corporate bond issuance, and intermediary activity. The Insurance Regulatory and Development Authority (IRDAI) collects data on life and non-life insurance premiums, penetration metrics, and claims ratios. The Pension Fund Regulatory and Development Authority (PFRDA) oversees statistics on National Pension System subscribers and asset allocation. The Association of Mutual Funds in India (AMFI) publishes assets under management and flow data. The National Payments Corporation of India (NPCI) and RBI jointly maintain statistics on digital payments—UPI, IMPS, NEFT, and RTGS transaction volumes and values. The National Bank for Agriculture and Rural Development (NABARD) compiles agricultural credit statistics.

This distributional model reflects an important principle: the institution most closely engaged with a particular financial domain possesses both the regulatory proximity and technical capability to extract meaningful data from market participants. The RBI’s access



Source: Mohan and Ray (2018)

Figure 1: Major constituents of Indian financial system

to supervisory information on bank capital adequacy and liquidity positions, for instance, is not merely statistical convenience but a regulatory imperative. Similarly, SEBI's direct relationship with market intermediaries and exchanges furnishes data that could not be reliably reconstructed from secondary sources¹.

Interestingly, various regulators operate within their own taxonomic frameworks, report at frequencies suited to their particular mandates, and deploy definitions that reflect sectoral custom rather than economy-wide standardization. The RBI publishes its Handbook of Statistics on the Indian Economy and Statistical Tables Relating to Banks in India at frequencies that have evolved with technological change, while the capital markets regulator operates on different publication schedules. Insurance and pension statistics follow their own rhythms. The data format, the compilation procedure, and the presentation often vary.

3. Data collection architecture

The methods through which financial data are extracted and processed reflect the regulatory and market infrastructure of modern finance. Statutory regulatory returns—mandated filings that market participants submit to regulators at specified frequencies—form the foundation. Banks file returns on deposits, advances, investments, and asset quality classifications. NBFCs report on the sectoral composition of their credit deployment. Insurance companies submit premiums and claims data. These returns are not voluntary or survey-

¹See Schipke *et al.* (2023) for some recent issues on Indian financial sector.

based; they are statutory obligations backed by regulatory authority.

Supervisory reporting systems layer additional detail over these mandated returns. The RBI's enhanced supervisory frameworks require banks to submit granular data on capital adequacy, liquidity coverage ratios, and net stable funding ratios at frequencies that have gradually shortened. Market infrastructure—the stock exchanges themselves, the depositories that hold securities in electronic form, the payment system operators—generate transaction-level data as a byproduct of settlement operations. This data, once the exclusive preserve of regulators, increasingly enters the public statistical domain, though usually in aggregated form.

Survey-based data and administrative records supplement the regulatory returns. The RBI conducts periodic surveys of household and business financial conditions, such as, Consumer Confidence Survey; Inflation Expectations Survey of Households; Survey of Professional Forecasters; Order Books, Inventories and Capacity Utilisation Survey; Industrial Outlook Survey; or Foreign Liabilities and Assets (FLA) Survey – to name a few. The Department of Financial Services tracks credit flows to priority sectors. Regulatory archives contain data on regulatory actions, supervisory corrections, and institutional changes. The challenge in this landscape is not data scarcity but rather the fragmentation of data sources, the heterogeneity of collection frequencies and methodologies, and the administrative burden this heterogeneity imposes on reporting institutions.

Timely data requires reporting at frequencies closer to real time. The days of annual or even quarterly data releases have given way to demands for monthly series, and in some cases, weekly or daily aggregations. The RBI publishes weekly data on monetary aggregates, reserve money, and foreign exchange reserves. The National Payments Council of India (NPCI) releases daily payment system statistics. Data released with a short lag from the reporting period are provisional and subject to revision as fuller information accumulates. A statistic deemed preliminary when released may be substantially revised when more complete information becomes available. The policy-maker or analyst who relies on the preliminary release risks basing decisions on data that diverges significantly from the later-revised figure.

Another challenge is the demands of international standards. Since finance is essentially global, the demands of financial statistics need to be consistent with various global standards. Annex 1 gives a description of the data collection architecture in India and various global standards.

4. The challenges being faced by financial statisticians

4.1. Financial innovation

Perhaps no force reshapes financial statistics more than financial innovation. The expansion of fintech, digital lending platforms, and cryptocurrency-based assets has created financial instruments and market mechanisms that sit uneasily within the taxonomic categories devised in earlier eras. When the RBI classifies financial institutions into scheduled commercial banks, cooperative banks, and non-banking financial companies, this taxonomy rests on historical distinctions that turn on licensing, deposit-taking authority, and regulatory jurisdiction. Yet a digital lending platform that extends credit through a technological interface, perhaps without taking deposits in the traditional sense, may not fit neatly into

these categories.

The challenge is not merely one of classification, but of preserving historical comparability while capturing the economic substance of new financial forms. If a peer-to-peer lending platform facilitates credit that would previously have been extended through a traditional bank, the growth in P2P lending may represent a genuine expansion of credit or merely a shift in the channel through which credit flows. Statistical frameworks that treat these channels identically will misrepresent the dynamics of financial activity. Yet frameworks that create new categories for each innovation risk fragmenting the time series and complicating longitudinal analysis.

Illustratively, cryptocurrency and distributed ledger technology present an even more acute variant of this problem. The Central Bank has legitimately articulated reservations about permitting unrestricted crypto-asset trading within India, citing monetary policy concerns and financial stability risks. Yet the emergence of stablecoins, digital rupee initiatives, and cross-border cryptocurrency transactions creates a parallel financial ecosystem that exists outside traditional regulatory perimeters and thus outside conventional financial statistics. The lack of data on cryptocurrency holdings, transaction values, and cross-border flows represents not merely a statistical gap but could turn out to be a potential blind spot in financial stability monitoring.

4.2. Shadow banking, NBFCs and institutional fragmentation

The rise of non-banking financial companies and alternative investment funds represents a conscious policy choice to broaden financial intermediation beyond the commercial banking system. The NBFC sector in India has grown substantially, encompassing housing finance companies, microfinance institutions, infrastructure finance entities, and other specialized lenders. Yet the regulatory boundary between what constitutes a reportable NBFC and what remains outside statistical compilation remains porous. Informal lenders, unregistered microfinance entities, and hybrid organizational forms operate in the interstices of regulation, extending credit without systematic reporting to any central authority.

This institutional fragmentation extends beyond the informal sector into the formal regulatory domain itself. A household's financial relationships may span multiple regulators' domains: deposits with a bank (RBI's purview), equity holdings (SEBI), insurance policies (IRDAI), pension contributions (PFRDA), and credit from an NBFC or financial technology platform (potentially outside systematic reporting). Yet financial statistics rarely aggregate across these domains to present a coherent picture of household financial position. The RBI collects data on bank credit to households; the SEBI publishes data on household equity investments; AMFI reports on mutual fund ownership; but no single statistical source synthesizes this fragmented information into a household balance sheet framework.

4.3. Household and MSME data gaps

The scarcity of high-frequency data on household balance sheets and financial conditions ranks among the persistent weaknesses in India's financial statistics. The RBI conducts the Surveys on Professional Forecasters and Consumer Confidence, providing periodic snapshots of expectations and perceived economic conditions. Yet systematic data on household

financial assets, liabilities, income sources, and wealth distribution remain fragmentary. This gap is particularly consequential because, in India's developing financial system, the household sector serves as a significant credit counterparty—both as a borrower of consumer credit and as a provider of non-institutional credit through informal channels.

The informal financial system remains quantitatively significant in India, particularly in rural areas and among lower-income households. The RBI estimates that informal credit continues to finance a substantial portion of consumption and investment by small and marginal farmers, non-farm entrepreneurs, and other economic agents outside the formal banking system. Yet this credit appears only obliquely in financial statistics, typically as a residual in household saving and borrowing calculations. The absence of systematic data on informal credit flows obscures the true trajectory of financial intermediation and introduces measurement error into assessments of monetary policy transmission.

Micro, small, and medium enterprises (MSMEs) face similar statistical invisibility. While the Reserve Bank collects data on credit flows to the MSME sector from formal lenders, information on balance sheet quality, default rates, and sectoral stress within MSMEs relies heavily on credit registry data from banks. Yet a significant portion of MSME financing occurs through trade credit, retained earnings, and informal borrowing mechanisms. The absence of comprehensive MSME balance sheet data—comparable to what regulatory supervisors maintain for banks and insurance companies—means that assessments of MSME financial health rest on partial information.

4.4. The granularity-privacy tension

As financial statistics have matured and data systems have become more sophisticated, both regulators and researchers have recognized the analytic power of granular, transaction-level data. Rather than analyzing aggregate credit flows to an economic sector, one can examine individual firm credit relationships, allowing for much richer analysis of credit allocation, concentration, and the transmission of monetary policy. Similarly, household-level data permit the analysis of distributional consequences of financial policy in ways that aggregate household statistics cannot.

Yet the movement toward granularity collides with legitimate concerns over privacy and data protection. The Digital Personal Data Protection (DPDP) Act, passed in 2023 with detailed rules notified in late 2025, imposes requirements on data collectors and processors regarding purpose limitation, consent, and security safeguards. The compliance framework specifies that personal data can be collected only for stated purposes and cannot subsequently be used for materially different purposes without fresh consent. The implications for financial statistics are substantial. A dataset initially collected for banking supervision may not be redeployed for monetary policy analysis or research without navigating the DPDP framework.

The DPDP Act establishes a significant compliance date of May 13, 2027, by which regulated entities must redesign their data collection, processing, and storage architectures to align with the act's requirements. The penalties for non-compliance are severe—fines up to ₹250 crore for serious breaches such as inadequate security safeguards or mishandling of children's data. These penalties create powerful incentives for institutions to err on the side of restricting data access rather than risking regulatory action. From a statistical perspective, the DPDP framework raises difficult questions. If anonymization is achieved through

aggregation, then granular data that permits identification of individual entities cannot be shared without consent. Yet consent frameworks that apply retroactively to data collected under earlier regimes create practical complications. A bank that has collected transaction data from millions of customers under the understanding that such data would be available for regulatory purposes may face substantial implementation challenges if it must now obtain fresh consent for each new analytic use.

The challenge becomes particularly acute in the context of cross-border data flows. The DPDP Act permits transfers of personal data to countries whose privacy frameworks the government approves, but the mechanism for whitelisting or blacklisting countries has not been fully clarified. International organizations, standard-setting bodies, and peer central banks who have historically relied on accessing Indian financial statistics may encounter new restrictions.

4.5. Cross-border integration

Financial integration with international capital markets has deepened over the past two decades. Foreign portfolio investors account for a significant share of trading on Indian stock exchanges. Foreign institutional investors hold substantial equity positions in Indian banks and corporations. Indian entities undertake cross-border borrowing in international debt markets. Correspondingly, the measurement of India's financial system and its stability cannot ignore the cross-border dimension.

Yet the standardization of financial data across borders remains incomplete. Different jurisdictions classify financial instruments differently, employ distinct definitional frameworks for capital adequacy, and report at different frequencies. The Basel Committee on Banking Supervision has promulgated common standards that major central banks implement, including India through the RBI. Yet implementation details—the precise definition of eligible capital, the risk weightings applied to different asset classes, the treatments of specific regulatory issues—vary across jurisdictions in ways that can complicate international comparisons.

4.6. Green finance

Additionally, the growing policy emphasis on climate-related risks and sustainable finance has introduced new dimensions to financial data collection for which standardized global methodologies do not yet exist. The Task Force on Climate-related Financial Disclosures (TCFD) has developed a framework for corporate climate risk disclosures, but adoption remains voluntary and variable. The RBI has issued guidance on climate risk management within the banking system and begun collecting data on climate-related lending, but this data collection is still nascent. The classification of "green" financing activities—renewable energy projects, energy efficiency investments, sustainable agriculture—lacks a universally accepted taxonomy despite ongoing international efforts toward standardization.

4.7. Artificial intelligence (AI) and alternative data

The rise of machine learning and large language models has begun to reshape the possibilities and risks in financial statistics. Regulators and academic researchers increasingly

deploy alternative data sources—transaction-level payment information, high-frequency market microstructure data, satellite imagery, social media sentiment—to construct financial indicators that complement or partially substitute for traditional statistical series. These alternative approaches can sometimes yield insights or identify economic phenomena faster than conventional data collection.

The integration of AI-driven statistical methods introduces substantial new challenges. Large language models, by their nature, operate as “black boxes” in which the relationship between inputs and outputs remains opaque even to their developers. A model trained on historical financial data may embed biases present in that data, propagating those biases into the statistics derived from its outputs. The phenomenon of “hallucinations” in LLMs—the generation of plausible-sounding but factually incorrect information—presents particular risks when applied to financial data where accuracy is critical.

The ethical implications warrant serious consideration. The deployment of machine learning algorithms in financial decision-making raises questions about the social acceptability and distributional consequences of automated decision-making. If credit allocation increasingly occurs through algorithmic processes rather than human judgment, and if those algorithms exhibit systematic bias, the result may be discrimination in credit availability that violates both legal norms and principles of financial inclusion. Transparency concerns do remain. The intellectual property claims that attach to proprietary algorithms, the commercial incentives to maintain algorithmic secrecy, and the technical complexity of explaining machine learning decisions all work against the kind of transparent, reproducible statistical practice that has been the aspiration of the statistical profession.

5. In lieu of conclusions: Toward stronger coordination and modern frameworks

Addressing these challenges requires movement on multiple fronts.

First, the various regulatory agencies compiling financial statistics would benefit from more systematic coordination on definitional frameworks and classification standards. While the RBI’s role as primary central bank for much of financial statistics provides a natural convening function, the growth of the NBFC sector, the increasing significance of capital markets, and the rise of fintech platforms mean that SEBI, IRDAI, PFRDA, and other regulators are now collecting material financial data. Harmonizing definitions, establishing common data standards, and coordinating on publication frequencies would reduce burden on reporting institutions and improve consistency across statistical domains.

Second, the architecture of financial data governance requires examination in light of the Digital Personal Data Protection Act and the legitimate concerns over privacy and cybersecurity. The RBI and other regulators must design data governance frameworks that balance the need for granular data—essential for effective financial stability monitoring and monetary policy—with the protection of personal information and the security of sensitive financial data. This might involve distinguishing between data required for regulatory purposes and data available for research or broader dissemination, establishing clear protocols for consent and data sharing, and investing in cybersecurity infrastructure proportional to the sensitivity of the data maintained.

Financial statistics matter because financial institutions and markets have become central to resource allocation, risk management, and monetary policy transmission in the Indian economy. Yet the statistical measurement of finance remains a work substantially in progress, populated by sophisticated institutional arrangements alongside notable gaps and inconsistencies.

The challenges identified in this talk —the revisionist timeline of data releases, the arrival of financial innovations that outpace statistical categories, the growth of non-banking finance, data gaps on households and small enterprises, the friction between privacy protection and statistical granularity, the complexities of cross-border financial integration, and the risks embedded in the application of artificial intelligence to financial measurement—are not aberrations in an otherwise smoothly functioning system. Rather, they reflect the inherent difficulty of measuring a complex, dynamic, and increasingly sophisticated financial system while balancing competing imperatives: timely data for policy-making, accuracy and stability in statistical series, protection of sensitive financial information, burden management on reporting institutions, and coordination across multiple regulatory agencies.

For statisticians engaged in financial measurement, the task ahead is both intellectually demanding and institutionally consequential. Robust financial statistics are vital not merely as technical artifacts but as essential infrastructure for financial stability, effective monetary policy, and the efficient allocation of capital in an economy increasingly dependent on sophisticated financial intermediation. The coming years will require greater coordination, ongoing investment in modern data infrastructure, and continued attention to the methodological challenges that arise as finance evolves.

Professor Ronald Coase, Nobel Laureate in Economics in 1991, reportedly said, “If you Torture the data enough, it will confess to anything.” The key task of the financial statistician is to get the confessions of the data without any torture. And that is a difficult task.

References

- Mohan, R., and Ray, P. (2018). Indian financial sector: Structure, trends and turns. In Ulrich Volz, Peter J. Morgan and Naoyuki Yoshino (ed.s): *Routledge Handbook of Banking and Finance in Asia*, London, Routledge.
- Schipke, A., Jarkko, T., Nada, C., and Gulde, A. M. (2023). *India’s Financial System: Building the Foundation for Strong and Sustainable Growth*. International Monetary Fund, Washinton DC.

ANNEXURE

Primary official sources on financial statistics

Reserve Bank of India:

- Handbook of Statistics on the Indian Economy (annual and updates)
- Statistical Tables Relating to Banks in India (published series)

- Financial Stability Report (half-yearly)
- RBI Bulletin (monthly; current data and policy analysis)
- RBI's reports on Monetary Policy (quarterly statements and supporting documents)
- Report on Trends and Progress of Banking in India (annual)

SEBI (Securities and Exchange Board of India):

- Annual reports and data on capital markets statistics
- Stock exchange handbooks and periodic market reports from NSE/BSE

IRDAI (Insurance Regulatory and Development Authority):

- Annual reports on insurance sector statistics
- Periodic industry reports on life and non-life insurance

PFRDA (Pension Fund Regulatory and Development Authority):

- Reports on NPS subscriber base and asset allocation
- Periodic disclosures on pension sector statistics

AMFI (Association of Mutual Funds in India):

- Monthly publications on assets under management
- Industry statistics and investor category data

NPCI (National Payments Corporation of India):

- Daily and periodic payment system statistics
- Reports on UPI, IMPS, NEFT, RTGS transaction volumes

NABARD (National Bank for Agriculture and Rural Development):

- Agricultural credit and rural finance statistics

Government of India:

- Economic Survey (annual; Statistical Appendix is particularly relevant)
- Department of Financial Services Annual Reports
- Union Budget documents

Relevant International Standards and Bodies

- Basel Committee on Banking Supervision publications (for capital adequacy and supervisory standards)
- Financial Stability Board reports on shadow banking and non-bank financial intermediation
- Task Force on Climate-related Financial Disclosures (TCFD) framework (for climate-related financial risk measurement)
- IMF and World Bank on financial sector measurement and statistics
- BIS (Bank for International Settlements)